Goals:

Distribute CLAS12 software
Submit simulation / reconstruction jobs @ JLab or @ Offsite resources
Reproducibility

Ongoing Work:

Secure offsite resources
Seamless submission to JLab’s or other farms

Use of Offsite resources

What's currently available

1. Open Science Grid (OSG)
2. MIT computing farms (Tier2, Tier3, possibly more)
3. Worldwide LHC Computing Grid (WLCG), for example INFN
4. UK Iris grid (also part of WLCG)
5. Glasgow University local computing cluster
6. GW local computing cluster (for GW users)
7. Jefferson Lab farm

This list is expanding. If your institution wants to contribute please contact us!

Use of Offsite resources

What are we shooting for?

- About 80 million CPU hours / year, corresponding to about 8,000 cores.
- Run single core jobs that include:
    1. Generators (or external LUND files),
    2. GEMC + baground from experimental data,
    3. Reconstruction.
    4. Filtering
- Simulation parameters defined in a Steering Card
- Submission framework common to all the computing resources
- Final output back here at Jlab or organized in other staging areas

Use of Offsite resources

- **Open Science Grid (OSG)**

We have an interactive node here at JLAB:

scosg16.jlab.org

JLab users can submit jobs to OSG using that machine.
This has been shown at the past 2 tutorials.

- We have an organization: org.CLAS12MC
- We currently run OPPORTUNISTICALLY, meaning we are paused / kicked out a node if something with higher priority comes along.
- Output is currently in the home directory (with plan to use /work or /volatile)

Available for us from scosg16. Snapshot at 4:10 on 6/17/2019

- 10826 Busy
- 665 Available

Use of Offsite resources

- **MIT computing farms (Tier2, Tier3, possibly more)**

The interactive nodes are at MIT

submit.mit.edu

The mechanism to submit jobs to these farms is identical to the OSG (HTCondor).
CLA12 users cannot directly submit jobs yet, will use the submission framework.

MIT group / graduate students working on:

Framework to prepare and dispatch simulation farm jobs to MIT, OSG and other off-site farms

- Bobby Johnston
- Sangbaek Lee
- Patrick Moran

Principal Research Scientist

- **Dr. Douglas Hasell**

Faculty

- **Or Hen**
- **Richard Milner**

Use of Offsite resources

- **UK Iris grid (also part of WLCG)**

Thanks to Amber / Peter Clarke (UK Grid for Particle Physics, GridPP management):

- UK grid is "officially" available to CLAS12 (JLab) users
- OSG to setup JLab node **scosg16** to (trivially) submit to GridPP. The outlook looks good as this was done for other organizations like CMS.
- Resources: David Ireland, Bryan McKinnon, Gareth Roy (Glasgow)

Other possible computing resources:

- Glasgow local computing farm
- Staging area of a few hundred Terabytes.

Use of Offsite resources

- **INFN (part of WLCG)**

The work on GridPP can be duplicated for the INFN resources

Request to Centro di Bologna, 2020-2023

500 Cores, equivalent to 5M CPU hours / year

50 TB

Use of Offsite resources

- **GW local computing cluster (for GW users)**

Glenn MacLachian, Computer administrator:

- Singularity, CVMFS installed

Giovanni Angelini: able to use on the farm:

- jeffersonlab/clas12simulations

Farm available to GW users only for now.

Use of Offsite resources

- **Jefferson Lab farm**

1. Singularity is installed.
2. CVMFS installation is in progress Wesley Moore (Computer Center) finalizing installation.

Amber Boehnlein, Bryan Hesse:

Jlab Computing Resources enthusiastic support to use containers

- Standardize the software tags: now it’s containers.
- Support for old legacy software trivial, no overhead.
- Transparent to OS changes.
- Running singularity is like opening a new shell.

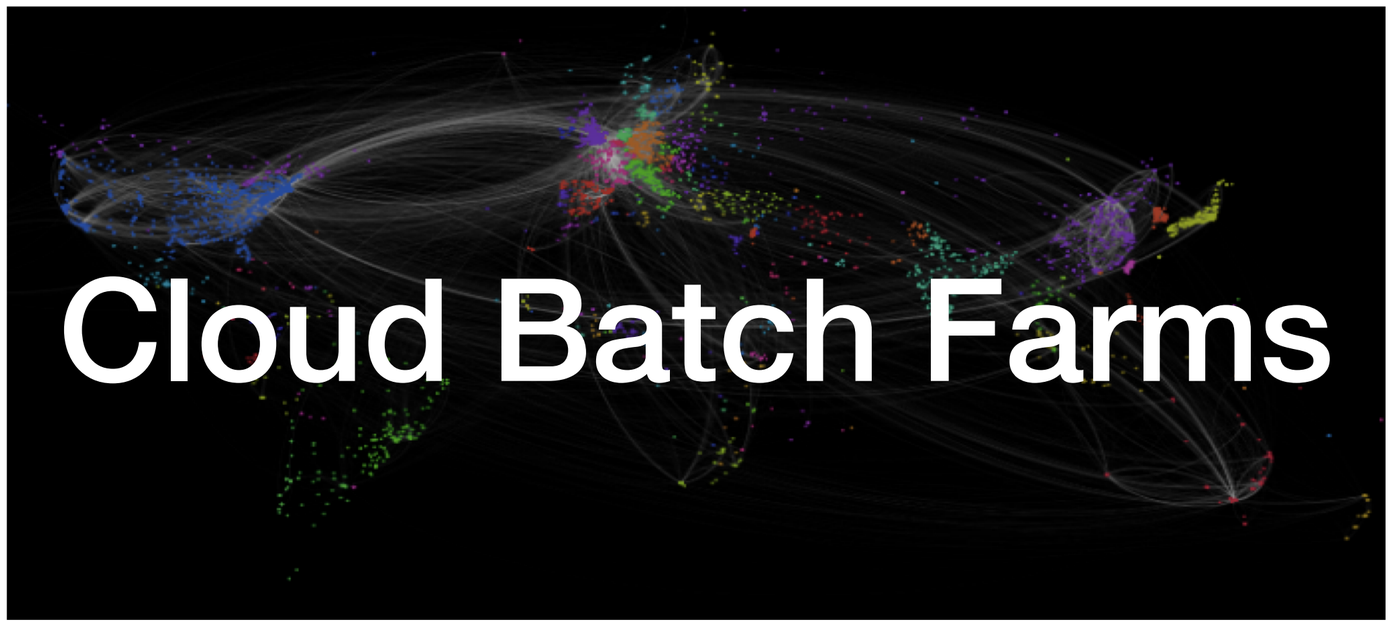
CLAS12 Offsite Farm Usage Vision

Steering card defines job.

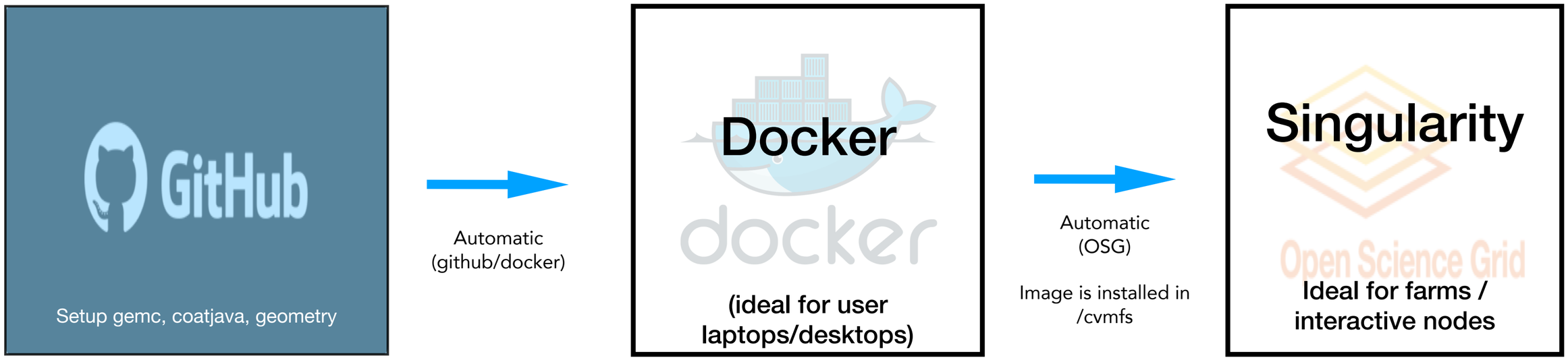
Choices include farm selection

Output Files
(JLab or other staging areas)

At JLab: Currently: 10 GB / s connection. FY2020: could be 100 GB connection.



GITHUB > CLAS12 DOCKER > Singularity connection (operational)



1. Changes to the Docker configuration trigger the docker image creation
2. The docker image creation trigger the singularity image creation
3. The singularity image is distributed to CVMFS

GITHUB > CLAS12 DOCKER > Singularity connection (operational)

What is CVMFS?

CernVM File System (CVM-FS): Files and directories are hosted on standard web servers and mounted (on demand) in the universal namespace /cvmfs.

We use CVMFS to distribute jeffersonlab/clas12simulations to offsite farms and the individual nodes:

```
/cvmfs/singularity.opensciencegrid.org/jeffersonlab/clas12simulations
```

Additionally we can use CVMFS for:

- Magnetic fields
- Geant4 data
- Standard background from experimental data
- Software libraries like ROOT

GITHUB > CLAS12 DOCKER > Singularity connection (operational)

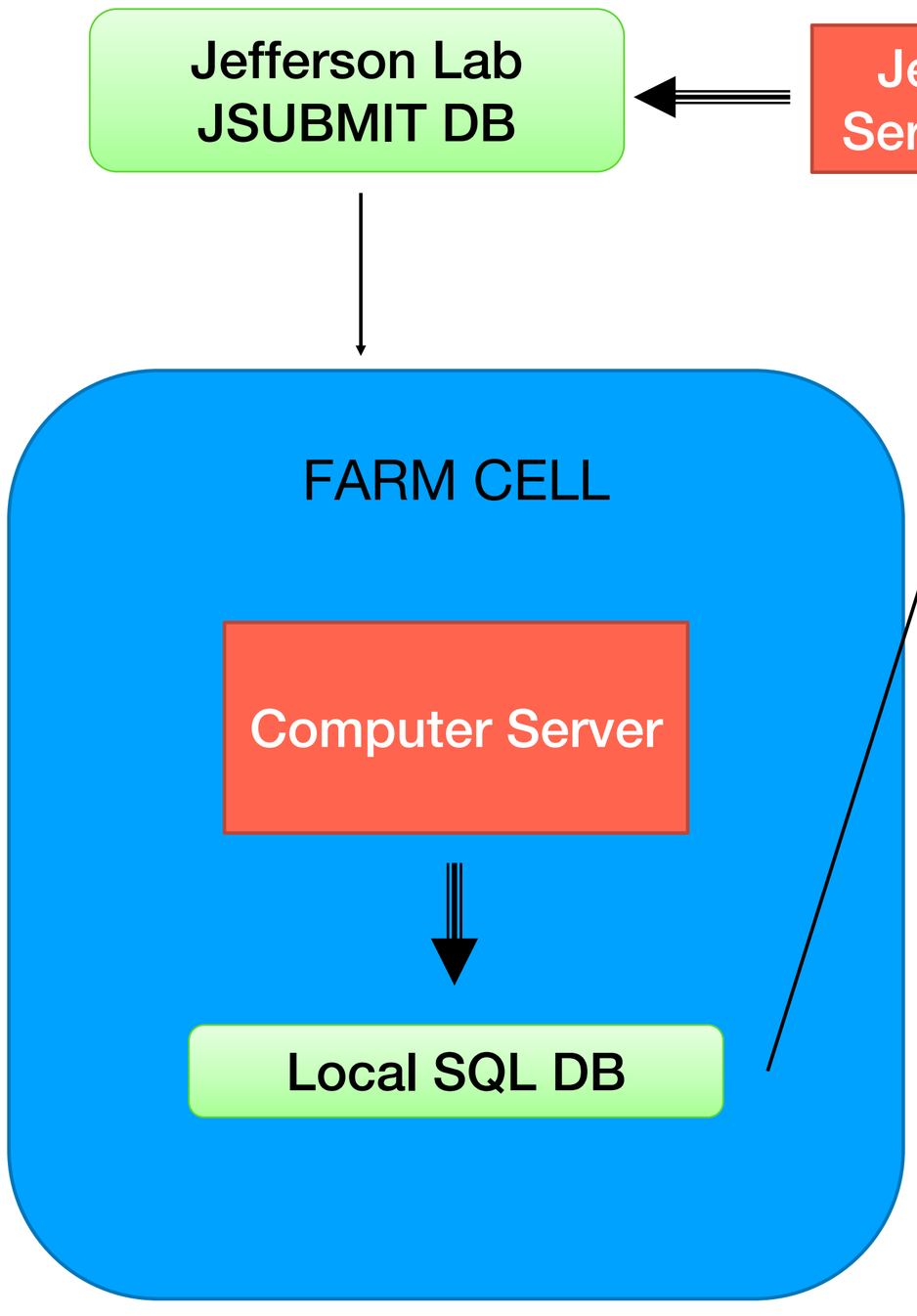
Notes:

- If your institution has or can install CVMFS, you'd have the CLAS12 software in your farm, for free.

- If in addition your institution has or can install singularity, you can run CLAS12 simulations and reconstruction, for free.

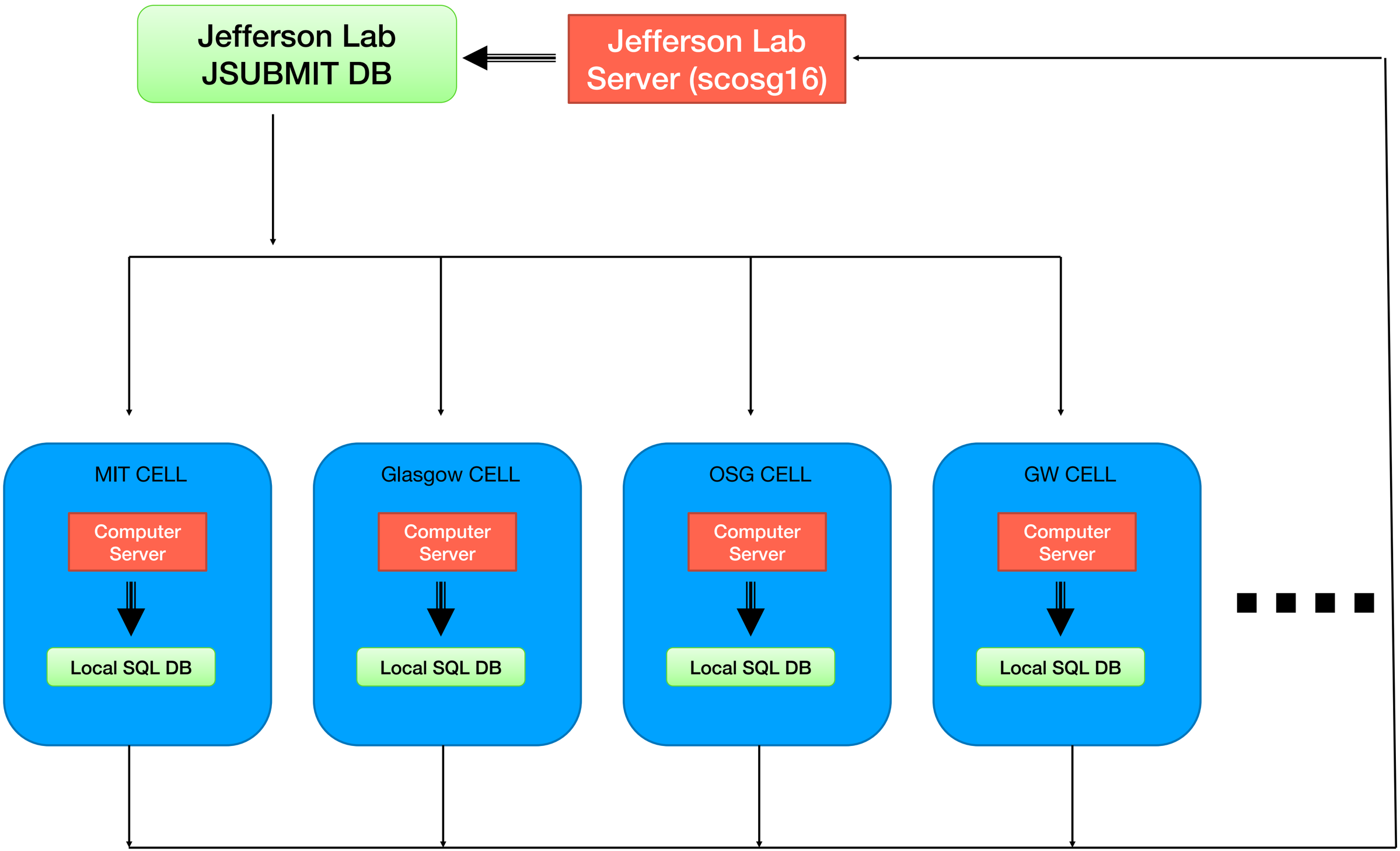
- If your institution can do both and has a farm, we could use it with minimum overhead. Contact us!

Farm “Cell”

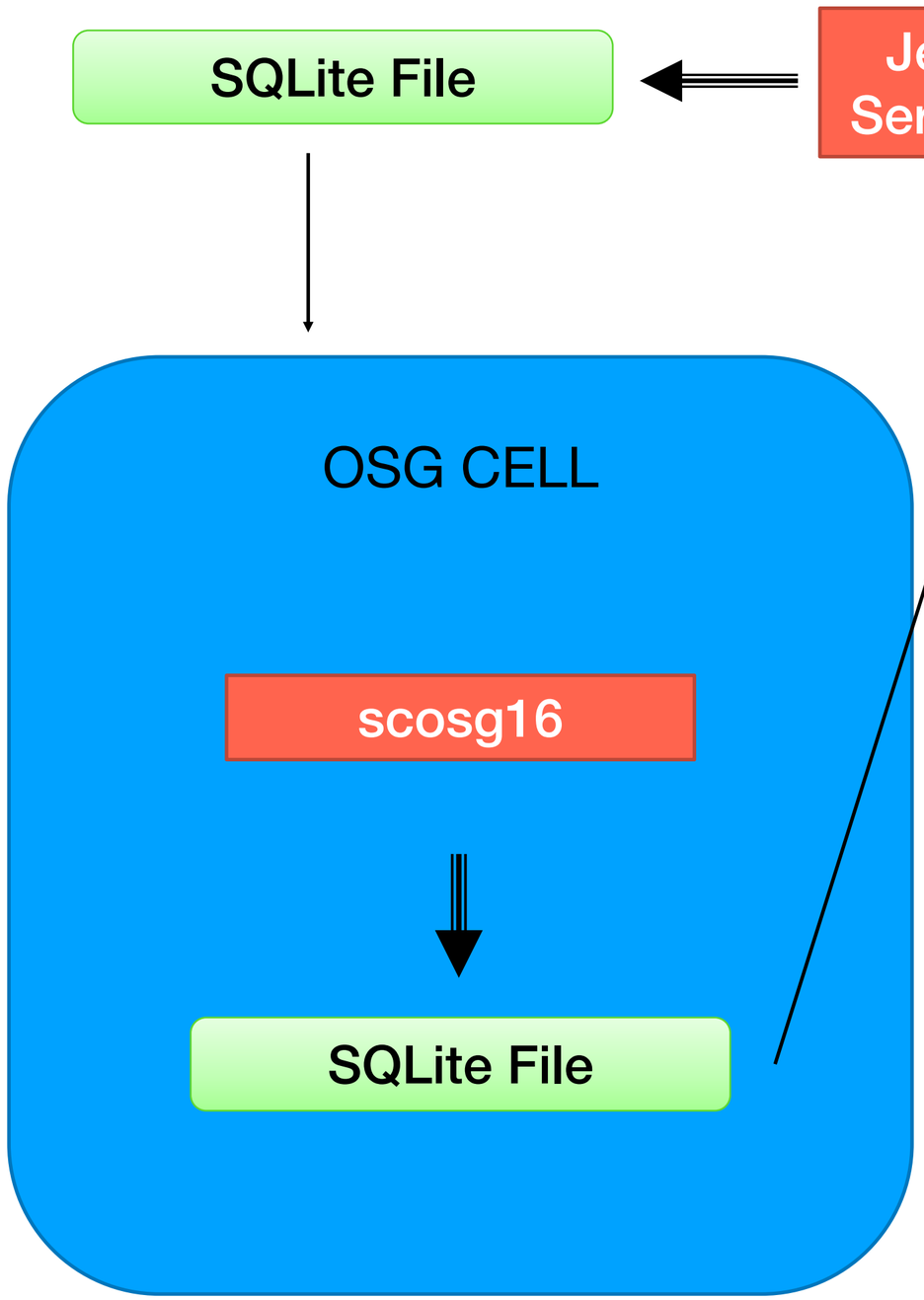


- Checks requests from JLab
- Submit jobs directed at it to its farm
- Communicate to JLab that the job is started / completed
- Send datafiles to JLab or staging area

Replicating Farm "Cells"

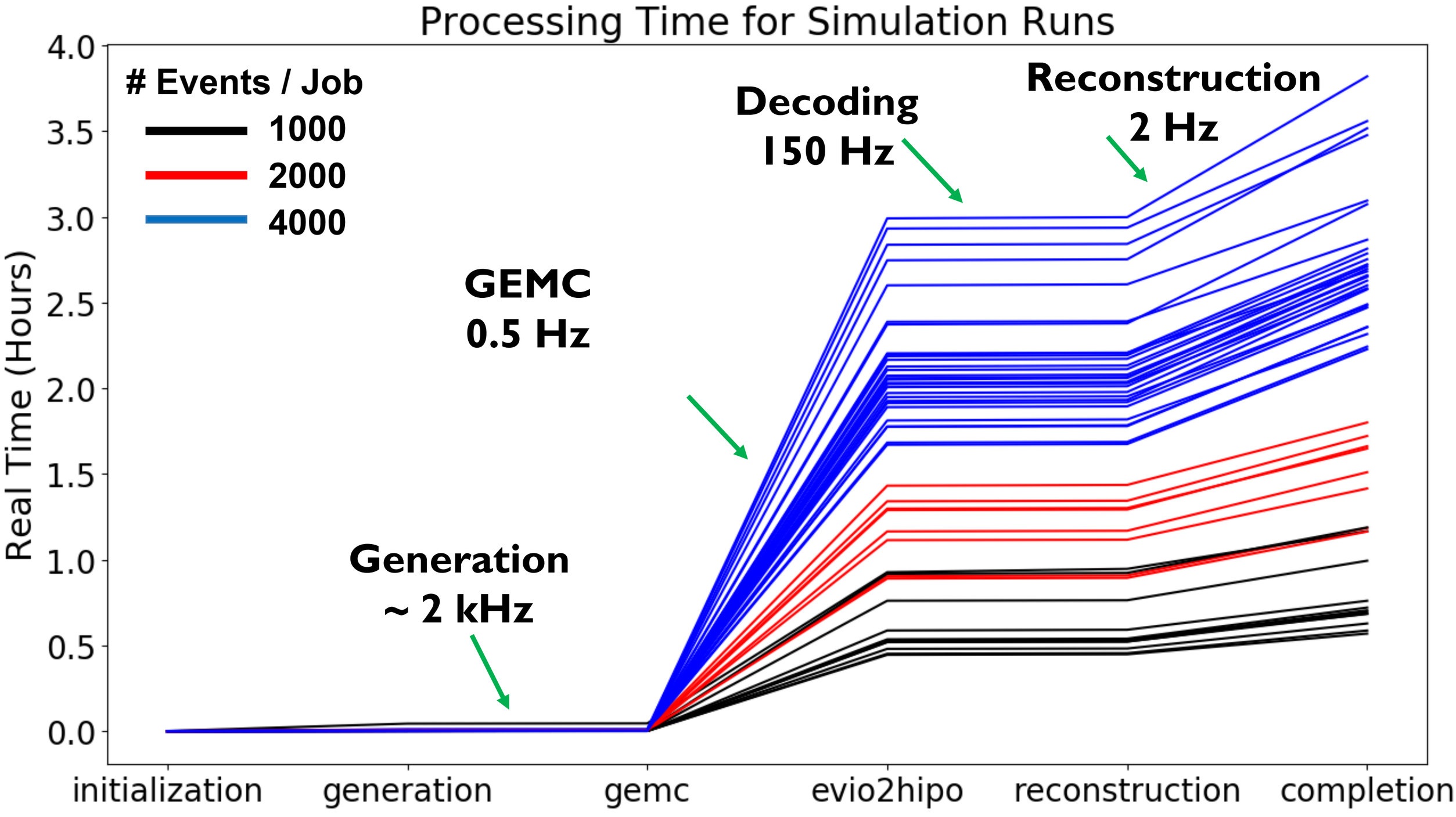


CLAS12 Offsite Farm Usage Status



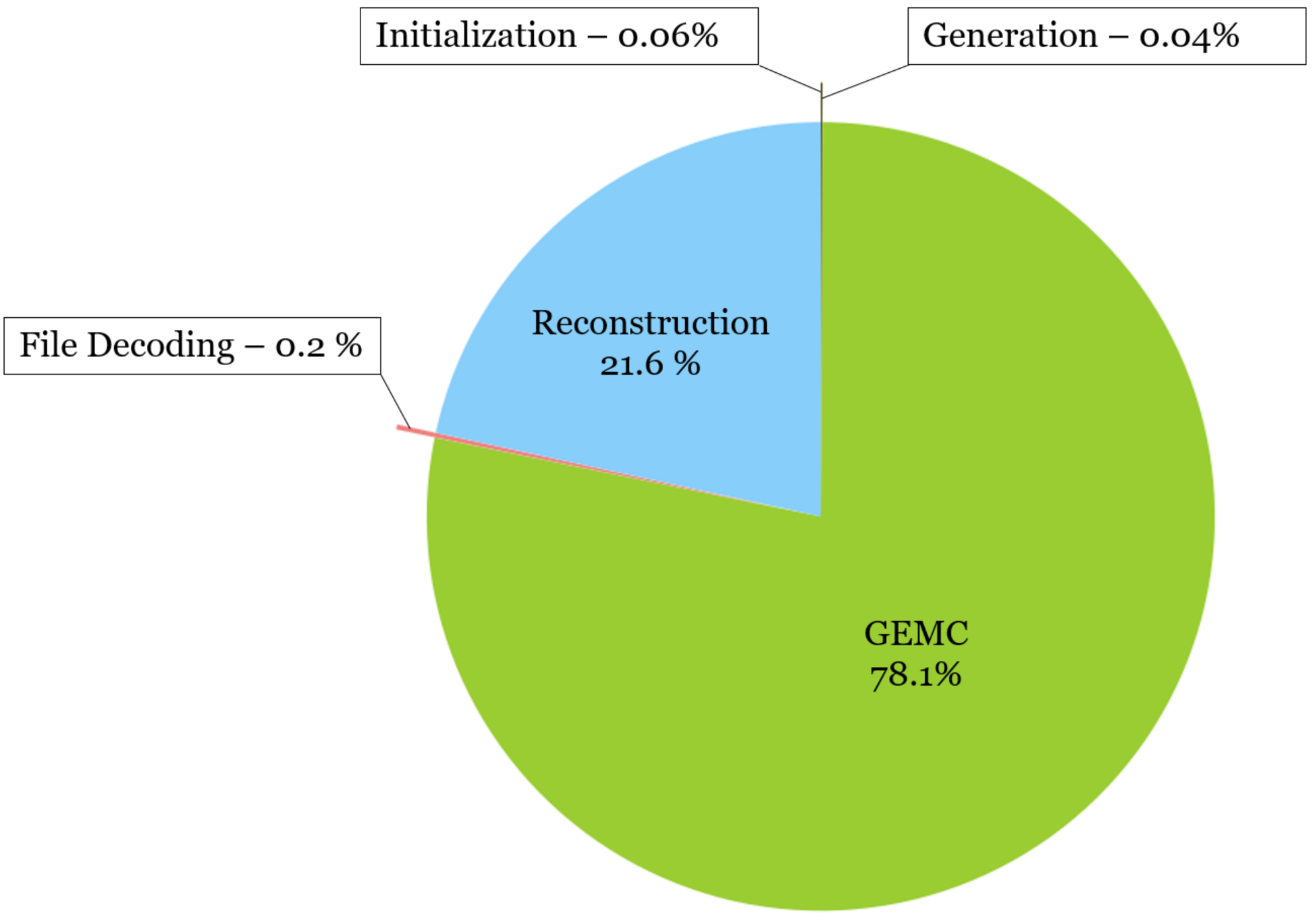
- Loop operational with SQLite
- jsubmit.jlab.org server setup
- Status: awaiting for security scans to read access from OSG nodes

Simulation Run Statistics



Simulation Run Statistics

Breakdown of Runtime



Steering Card

Example 1: 1,000,000 events to OSG with CLASDIS generator, run group A

```
project:  osg.clas12MC          # project name
group:  rgaDIS                  # project description
farm_name: OSG                  # farm pool
generator: clasdis              # one of clasdis, dvcs, disrad.
nevents:  10000                 # number of events each job
gcards: rga-fall2018            # experiment gcard
luminosity:  0                  # percent of 10^35 luminosity
tcurrent: -1.0                  # torus field scale
pcurrent: -1.0                  # solenoid field scale
cores_req: 1                    # number of cores to request from node
mem_req: 2                      # GB of RAM to request from node.
jobs: 100                       # number of jobs for each submission.
filter: yes                     # Filter the output so that
keepMCData: no                  # Keeps GEMC evio or hipo file
```

Steering Card

Example 2: run @ Mit Tier 2, run group A, user LUND files
Keep gemc hipo files

```
project:  osg.clas12MC                                  # project name
group:  rgaDIS                                          # project description
farm_name: MIT_Tier2                                    # farm pool
generator: https://userweb.jlab.org/~ungaro/lund/       # location of user lund files.
gcards: rga-fall2018                                    # experiment gcard
luminosity:  0                                          # percent of 10^35 luminosity
tcurrent: -1.0                                          # torus field scale
pcurrent: -1.0                                          # solenoid field scale
cores_req: 1                                            # number of cores to request from node
mem_req: 2                                              # GB of RAM to request from node.
filter: yes                                             # Filter the output so that
keepMCData: hipo                                        # Keeps GEMC evio or hipo file
```

There is exactly one job / file
Nicely Organized in subdirectories

Steering Card

Example 3: run @ Mit Tier 2,
10,000 events, 10 event / job
Compare two experimental setup, with physics background
Keep EVIO file

```
project:  osg.clas12MC                                  # project name
group:  rgaDIS                                          # project description
farm_name: MIT_Tier2                                    # farm pool
gcards: https://userweb.jlab.org/~ungaro/gcards/        # If online address,
                                                        # there will be one submission for
                                                        # each gcard at that address
nevents:  1000                                          # number of events each job
luminosity:  100                                        # percent of 10^35 luminosity
tcurrent: -1.0                                          # torus field scale
pcurrent: -1.0                                          # solenoid field scale
cores_req: 1                                            # number of cores to request from node
mem_req: 2                                              # GB of RAM to request from node.
jobs: 100                                               # number of jobs for each submission.
                                                        # This entry is ignored if lund files are used.
                                                        # In that case, theres is exactly one job / file
filter: yes                                             # Filter the output so that
                                                        # only reconstructed banks are saved
keepMCData: evio                                        # Keeps GEMC evio or hipo file
```

GEMC 4.3.1, Containers and Documentation

CLAS12 Simulation Software Distribution

home | examples

The CLAS12 software packages are distributed using docker images.

Quickstart: Docker

Use the following command to run the clas12 software image:

```
docker run -it --rm jeffersonlab/clas12simulations:iprod bash
```

Container content

The clas12simulations docker image contains:

- various generators
- gemc with the clas12 geometry
- CLARA
- Coatjava
- the *CLAS12_BIN*, *CLAS12_LIB*, *CLAS12_INC* dirs and environment variables

For the packages version check the tags page.

HowTos

- Batch mode
- Mounting your directories to the container
- Graphic mode (**browser**)
- Graphic mode (**vnc**)
- **Native interactive** mode (no opengl)
- Using Generators
- Convert GEMC evio output to hipo
- Reconstruction

Tags

- Production (currently 1.2)
- 1.1
- 1.0

CLAS12 Tags

Production:

- 4.3.0: **COATJAVA: 5.7.4, JLAB_VERSION: 2.3**:
  - Updated DC geometry using latest survey (May 18 Entry in DB)
  - Fixed bug that prevented material name from being displayed in the GUI
  - 3d cartesian field map support
  - new geant4 version: 10.4.p02
  - 51 micron tungsten shield (for bst) surrounding the target
  - calorimeters: reading ecal effective velocity from CCDB
  - change htcc time offset table to use the same used in reconstruction
  - Tony Forest: Added polarized target geometry/material and cad volume.

Frozen
Installed on farm
Installed on docker/ singularity
(offsite resources)

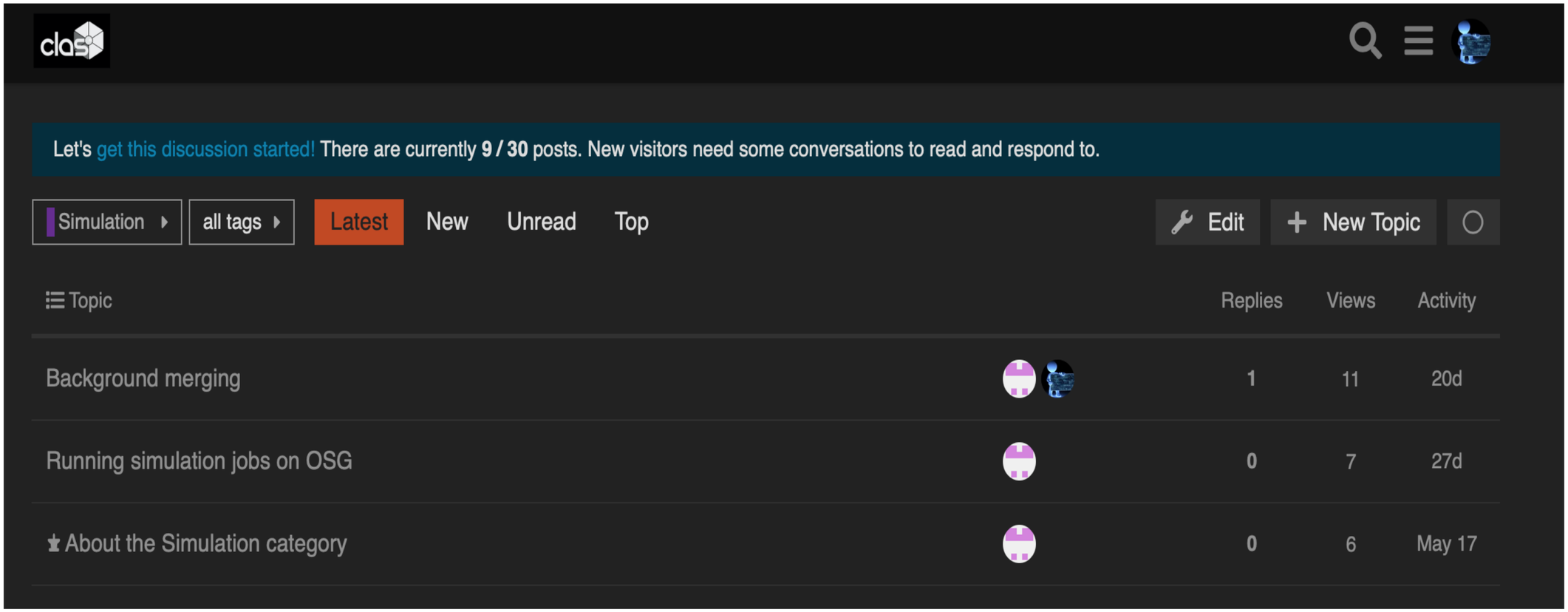
In development:

- 4.3.1 (By May 2019):
  - FTOF Time resolution updated based on data
  - Option SAVE_SELECTED, RERUN_SELECTED to save RNG state for certain particles, detector
  - Option SAVE_ALL_ANCESTORS to save complete particles hierarchy in output (evio2root also updated)
  - gcards for rg-a different run-periods
  - gcards for rg-b different run-periods
  - ec, pcal digitization removed obsolete constants
  - moved ftof shield in the correct position
  - Option written in JSON format
  - rga_fall2018 variations for: FTOF, EC, PCAL, CTOF geometry services
  - default variation for DC geometry service
  - ltcc variarions for different run periods ➡ SOON
  - target position added to ctof digitization shift

Installed on farm for testing
To be Installed on docker/ singularity
To be frozen after testing

Linked from CLAS12 Software Center

Discourse



Please post your simulation questions here!

Run Configurations to be released in new container (and we'll demo these gcards)

- Run group A Spring 2018:
  - Central detector shifted 19.4mm upstream
  - target (LH2) at (0, 0, -19.4) mm
  - HTCC shfted 9.5mm upstream
  - FT On configuration
  - FMT present
  - LTCC sectors: 2 (N2), 3 (N2), 5 (old C4F10), 6
  - Torus polarity: -1, 1, -0.75, 0.75
  - Solenoid polarity: -1
  - Beam Current: from 5 to 75 nA

rga-spring2018.gcard

- Run group A Fall 2018:
  - Central detector shifted 30 mm upstream
  - target (LH2) at (1.2, 1.1, -30) mm
  - HTCC shfted 19.5mm upstream
  - FT On configuration
  - FMT not present
  - LTCC sectors: 3 (50% C4F10), 5 (N2)
  - Torus polarity: -1, 1,
  - Solenoid polarity: -1
  - Beam Current: from 5 to 75 nA

rga-fall2018.gcard

Summary

Securing offsite resources:
MIT, GridPP, INFN, Glasgow, GW, Other WLCG

Seamless submission to JLab or other farms:
about to test JLab farm cell loop on MYSQL

Steering card mechanism allows to completely define the simulation parameters.

Docker image is being adopted as the CLAS12 software Distribution, even when you need to develop new geometry/algorithms

Friday Demo: more on docker usage

Robby Johnston, MIT

How to use the docker image to test your own gemc version.
How to use your own LUND files

A look at the Steering Card for off-site submissions